Title	5. 非Shannon情報量による情報源符号化(第9回『非平衡系の統計物理』シンポジウム,研究会報告)
Author(s)	山野, 拓也
Citation	物性研究 (2002), 77(5): 918-921
Issue Date	2002-02-20
URL	http://hdl.handle.net/2433/97170
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

非Shannon情報量による情報源符号化

山野　拓也
東京工業大学物性物理専攻

1　はじめに

現在、Tsallisによるボルツマンエントロピーを非加法的に拡張したエントロピーに基づく統計力学の形式が[1]注目を集めている。そこではパラメーターqを用いて$q \to 1$の極限でボルツマンエントロピーになるようにしている。ところで、ボルツマンによる統計力学的エントロピーと情報理論におけるシャノンによる情報エントロピー（シャノンエントロピー）は同じ形であることはよく知られている。そこで統計力学の非加法的拡張が行われているのであれば、情報理論についてもそのような拡張ができるのではないかと考えるのは自然であろう。本講演では従来の情報理論において成り立ついくつかの定理が、非加法的な情報量を基にする形式において拡張されることを報告した。情報理論においては、情報源から生じるメッセージをいかに効率よく符号化して通信路に乗せるかという問題と、一般にノイズのある通信路を経由してきた符号語をどのようにうまく復号して、情報源からの元のメッセージを推定するかという問題に分かれる。数学的には前者は情報源符号化定理、後者は通信路符号化定理と呼ばれる。本講演では前者の非加法的拡張について報告した。ここで言う拡張という意味は、Tsallis統計力学の場合と同様に$q \to 1$の極限でシャノンの情報理論を含むことである。

以降に導入するいくつかの情報理論的に重要な関係式と定理の具体的な証明については文献[2]を参照していただきたい。

2　情報量の定義とエントロピー

出発点として情報理論において基本的な量である情報量の定義からはじめる。シャノンが"通信における数学的理論"というオリジナル論文[3]のなかで採用している情報量$I(p)$は、ハートレイによる$I(p) \equiv -\ln p(x)$（$p(x)$は事象xが情報源$\mathcal{H}$から生じる確率、この場合単位はナット）である。これをTsallis統計力学で用いられるq-対数関数（$\ln_q x = (x^{1-q}-1)/(1-q)$）に置き換えて非加法的な情報量として、$I_q(p) \equiv -\ln_q p(x)$を採用することにする。これは$q \to 1$で$I(p)$になり、また情報量のもつ直感的に理解できる単調減少性をもつ点で非加法的な拡張として好ましいと思われる。この単調減少性とは、当然起こることが起きた時に、それを知っても得る情報量はゼロであるが、起こりそうもないことが起きた時には得る情報量は大きくなる事を表している。ただし$I_q(p)(q \neq 1)$の場合には$p(x)=0$で発散せず、上限値$1/(1-q)$があることに注意する。次に離散的な事象を生じる情報源のもつ情報エントロピーを定義する。情報理論における情報エントロピーとは、一事象当りの平均情報量として定義されている。ここでは通常の意味での平均のかわりにTsallis統計力学で用いられるいわゆるエスコート平均[4]によって平均をとる

ことにする。すなわち情報エントロピーは、

$$H_q(X) = \frac{\sum_{x\in\mathcal{H}} p^q(x) I_q(p)}{\sum_{x\in\mathcal{H}} p^q(x)} = \frac{1-\sum_{x\in\mathcal{H}} p^q(x)}{(q-1)\sum_{x\in\mathcal{H}} p^q(x)} \tag{1}$$

のように表される。この形は Tsallis エントロピーを因子 $\sum_{x\in\mathcal{H}} p^q(x)$ で割った形をしていることに注意したい。情報理論で現れるその他のエントロピーについても、条件付き情報量 $I_q(x \mid y)$、結合情報量 $I_q(x,y)$ をそれぞれ q-対数関数で定義して

$$H_q(Y \mid x) = \frac{1-\sum_{y\in\mathcal{Y}} p^q(y \mid x)}{(q-1)\sum_{y\in\mathcal{Y}} p^q(y \mid x)}, \quad H_q(X,Y) = \frac{1-\sum_{x\in\mathcal{H},y\in\mathcal{Y}} p^q(x,y)}{(q-1)\sum_{x\in\mathcal{H},y\in\mathcal{Y}} p^q(x,y)} \tag{2}$$

とする。ここで X や Y は事象の実現値 x、y の属性を示す。$\mathcal{Y}$ は y の属する情報源である。また、条件付きエントロピーに関して次の定義を導入する。

$$\left\langle \frac{1}{1+(q-1)H_q(Y \mid x)} \right\rangle_q^{(X)} \equiv \frac{1}{1+(q-1)H_q(Y \mid X)} \tag{3}$$

ここで $\langle\cdot\rangle^{(X)}$ は $p(x)$ についてエスコート平均をとることを意味する。このように準備すると X と Y の非加法的な結合エントロピーは次の関係式を満たすことがわかる。

$$H_q(X,Y) = H_q(X) + H_q(Y \mid X) + (q-1)H_q(X)H_q(Y \mid X). \tag{4}$$

X と Y が独立のときには $H_q(Y \mid X) = H_q(Y)$ が成り立つので、擬加法性（ただし今の場合,$q > 1$ と $q < 1$ とが Tsallis エントロピーの場合とは逆）が再現される。この関係式から Y の X に対するあいまい度、すなわち X の生起を条件とする Y の生起についての条件付きエントロピー $H_q(Y \mid X)$ は、もう一つの事象系 Z を使って

$$H_q(Y \mid X) = \frac{H_q(Y,Z \mid X) - H_q(Z \mid Y,X) + (q-1)\{H_q(X)H_q(Y,Z \mid X) - H_q(X,Y)H_q(Z \mid Y,X)\}}{1+(q-1)H_q(X)} \tag{5}$$

と表せる。これらの関係式を使うと事象 X_1 から X_n までの結合エントロピー $H_q(X_1, X_2, \cdots, X_n)$ には次のような階層的な構造があることが分かる。すなわち、

$$H_q(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} \left[1+(q-1)H_q(X_{i-1}, \cdots, X_1)\right] H_q(X_i \mid X_{i-1}, \cdots, X_1) \tag{6}$$

となり、n 個の事象の結合エントロピーを知るには、それ以前の $n-1$ 個の結合エントロピーを知る必要がある。ここまでは、情報エントロピー自体の性質である。

次に情報理論で重要な量である Kullback-Leibler(KL) の相対エントロピーを定義する。シャノンの情報理論における KL エントロピーが二つの事象に伴う情報量の差を一方の確率分布で平均したものとして定義されることを考慮して、非加法的な拡張の場合についても同様に非加法的情報量の差 $\Delta I_q \equiv I_q(p'(x)) - I_q(p(x))$ をエスコート平均したもので定義することにしよう。すると

$$D_q(p(x) \parallel p'(x)) \equiv \frac{\sum_{x\in\mathcal{H}} p^q(x)\Delta I_q}{\sum_{x\in\mathcal{H}} p^q(x)} = \frac{\sum_{x\in\mathcal{H}} p^q(x)\left(\ln_q p(x) - \ln_q p'(x)\right)}{\sum_{x\in\mathcal{H}} p^q(x)} \tag{7}$$

となって、これは $q>0$ のときに非負であることが証明できる。また $p(x)=p'(x)$ のときに $D_q(p(x)\parallel p'(x))=0$ となる。相互情報量についても形式的に $I_q(Y;X)\equiv H_q(Y)-H_q(Y\mid X)$ というように定義することができるが、シャノンの情報理論のように X と Y について対称とはならず（$I_q(Y;X)\neq I_q(X;Y)$）、別の定義を用いるべきか今のところ不明である [5, 6]。しかし情報理論で必要な $I_q(X;Y)$ の正値性を仮定すれば、結合エントロピーに対して次の不等式が成り立つことが証明できる。

$$H_q(X_1,X_2,\ldots,X_n)\leq\sum_{i=1}^{n}\left[1+(q-1)H_q(X_{i-1},\ldots,X_1)\right]H_q(X_i) \tag{8}$$

3 情報源符号化

前節までがセットアップである。この節では前節の準備により、シャノンの情報理論において無記憶情報源、あるいはマルコフ情報源に対して成り立つ、エントロピーと KL エントロピーの間の不等式を非加法的な場合にも拡張できることを述べる。
情報源のアルファベットの数が M である離散的な無記憶情報源から符号語 $X_i(i=1,\cdots M)$ の長さ（X_i を構成するシンボルの個数）が l_i であるような符号語が生起する場合、その符号語からなるメッセージが一意的に復号可能であるための必要十分条件は Kraft 不等式と呼ばれる

$$\sum_{i}^{M}D^{-l_i}\leq 1 \tag{9}$$

という関係式を満たす。ここで D は、アルファベットに属するシンボルの個数である（バイナリーの場合は $D=2$）。一般に情報源符号化においては符号語の長さをできるだけ短くなるような符号化を行いたい。したがって、平均の符号長（ここでも平均は l_i をエスコート平均することを意味する）に対して Kraft の不等式を拘束条件とする最適化問題を解くことになる。即ち考えるのは Lagrange の未定乗数 λ を使って、Kraft の不等式の等号を考えた汎関数

$$J=\frac{\sum_i p_i^q l_i}{\sum_i p_i^q}+\lambda\left(\sum_i D^{-l_i}\right) \tag{10}$$

を最小にするような場合である。これにより最適化符号長 l_i^* は $\log_D(\sum_i p_i^q)-q\log_D p_i$ と表される。しかし符号語の長さは整数のはずであるが、この l_i^* は必ずしも整数になるとは限らないので、整数に丸めた $l_i=\lceil\log_D(\sum_i p_i^q)-q\log_D p_i\rceil$ が実際の符号語の長さであると考える。ここで記号 $\lceil x\rceil$ は、x より大きい最小の整数を表す。この整数の符号長は

$$\log_D\left(\sum_i p_i^q\right)-q\log_D p_i\leq l_i<\log_D\left(\sum_i p_i^q\right)-q\log_D p_i+1 \tag{11}$$

という関係式を満たすので、$p_i^q/\sum_i p_i^q$ をかけて i について和をとると平均の符号長 $\langle L\rangle_q$ は

$$H_q(p)+D_q(p\parallel r)\leq\langle L\rangle_q<H_q(p)+D_q(p\parallel r)+1. \tag{12}$$

という関係を満たすことがわかる。ここで r は p と異なる確率分布で、p と異なる確率分布で情報源から生起すると一般化された KL エントロピーの分だけペナルティーがつくことを表している。q という添字がつくことを別にすればこれは、シャノンの情報理論で成

り立つ関係式と同じである。これは n 文字からなるメッセージの符号化についても一般化できる。すなわち n 文字の平均の符号長 $\langle l(X_1,\dots,X_n)\rangle_q$ は

$$H_q(X_1,\dots,X_n) \le \langle l(X_1,\dots,X_n)\rangle_q < H_q(X_1,\dots,X_n)+1 \tag{13}$$

を満たすので、一文字あたりの平均符号長 $\langle L_n\rangle_q = \langle l(X_1,\dots,X_n)\rangle_q/n$ については、

$$\frac{\sum_{i=1}^{n}\left[1+(q-1)H_q(X_{i-1},\dots,X_1)\right]H_q(X_i)}{n} \le \langle L_n\rangle_q < \frac{\sum_{i=1}^{n}\left[1+(q-1)H_q(X_{i-1},\dots,X_1)\right]H_q(X_i)}{n} + \frac{1}{n} \tag{14}$$

となる。ここで Eq.(8) を用いた。これが情報源符号化定理の情報量 $I_q(p)$ に基づく非加法的な拡張と考えられる。

4 まとめ

本講演では非加法的統計力学とのアナロジーを動機として非加法的な情報理論の可能性を形式的に調べて見た。冒頭の"はじめに"でも述べたように情報理論のもうひとつの柱である通信路符号化定理の拡張についてはこの講演では触れなかった。通信路符号化を考えるためには拡張された相互情報量をコンシステントに定義する必要があり、現在までのところ満足のいく定義はない。しかしいくつかの試みはある [5, 6]。この形式の実際の工学的な応用可能性については今後の研究の発展を期待したいところである。

参考文献

[1] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).

[2] T.Yamnao, "Information theory based on nonadditive information content", Phys.Rev.E **63** 46105 (2001).

[3] C.E.Shannon, Bell Syst. Tech. J. **27**, 379 (1948); *ibid* 623 (1948); C.E. Shannon and W.Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana,1963).

[4] C. Tsallis, R.S. Mendes and A.R. Plastino, Physica A **261**, 534 (1998).

[5] T.Yamano, eprint, cond-mat/0102322 ; T.Yamano, "Comment on the generalized mutual information", in preparation.

[6] Private communication with Prof. A.K. Rajagopal.